Cognition 262 (2025) 106162

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Full Length Article

Environmental context scaffolds children's semantic representation of novel words

Elise Breitfeld [*], Jenny R. Saffran

Department of Psychology, University of Wisconsin-Madison, USA

ARTICLE INFO

Keywords:
Word learning
Nouns
Meanings
Environmental context
Verb-mediated prediction
Extra-linguistic cues

ABSTRACT

In their everyday lives, children encounter words and objects in meaningful environments; for example, food-related words and objects tend to appear in the kitchen. The current study incorporated this aspect of children's naturalistic word learning experience into an experimental paradigm designed to examine whether environmental context impacts the meanings children ascribe to novel nouns. Preschoolers (36–48 months, $N = 46$) heard labels for novel objects embedded in images of natural scenes (in the kitchen or outdoors). They were then tested using a verb-mediated prediction paradigm. Children heard sentences where one of the novel labels was preceded by either a neutral verb ("see" or "find") or a context-related verb ("eat" or "throw") while viewing object pairs. The results showed that children used the context-related verbs to anticipate the target noun. This pattern of results suggests that children encoded environmental context information during word learning and used it to inform their representations of the meanings of novel words.

1. Introduction

Young children rapidly become expert word learners, capable of extracting useful information from their complex surroundings. The cluttered environments within which word learning occurs are often viewed as a hinderance to children's successful word learning: there are many objects present, and many activities occurring beyond explicitly learning-focused interactions, that could obscure word-referent learning (e.g., Quine, 1960). A large body of word learning literature has explored how children contend with this challenge, deemed the mapping problem, to successfully connect nouns and referents (see Wojcik et al., 2022). As a result of this focus on the referential ambiguity inherent in children's word learning environments, the possibility that these environments could contribute positively to the word learning process remains largely unexplored. The current study examines how word learning environments might help scaffold one particular aspect of word learning: building semantic knowledge beyond simple object-label mappings.

Critically, the meanings of nouns encompass much more than just what their referents look like. For example, while the word *banana* refers to objects with particular physical features (i.e., yellow, crescent-shaped), other aspects of its meaning include that it can be eaten, it grows on trees, and it is associated with monkeys. By age two, toddlers begin to demonstrate semantic knowledge about nouns that goes beyond the visual features of their referents (e.g., Arias-Trejo & Plunkett, 2013; Mani & Huettig, 2012; Willits et al., 2013). However, exactly how children acquire this knowledge and the role that the environment might play in the process remain open questions.

One key source of semantic knowledge is linguistic input. Children may be able to learn aspects of word meanings that extend beyond the physical properties of their referents from both explicit input, like generic statements, and implicit input, like co-occurrence statistics. For example, caregivers might say something like "Bananas are yummy." This kind of statement is known as generic language, which describes properties of a noun category (e.g., "Bananas are yummy") rather than properties of a particular individual referent (e.g., "This banana is yummy"). Indeed, there is evidence that, by preschool-age, children can recognize generic language and use these statements to learn non-observable properties of novel noun categories (e.g., "Blicks drink milk;" Gelman et al., 2010; Gelman & Raman, 2003; Graham et al., 2011). However, there is somewhat mixed evidence regarding how frequent this generic language is in children's input—with estimates ranging from only about 1 % of utterances up to around 10 % (Gelman et al., 2014; Pappas & Gelman, 1998; Wei et al., 2022). Other, more implicit, elements of language input may also contribute to building semantic knowledge (Wilson et al., 2023). For example, toddlers can

* Corresponding author at: 1202 W. Johnson St, Madison, WI 53706, USA.
E-mail address: breitfeld@wisc.edu (E. Breitfeld).

https://doi.org/10.1016/j.cognition.2025.106162
Received 6 August 2024; Received in revised form 17 April 2025; Accepted 24 April 2025
Available online 30 April 2025

identify relationships between semantically-related nouns on the basis of direct co-occurrences in language (e.g., "apple" and "banana" often occur close to one another in speech) and shared co-occurrences (e.g., "apple" and "banana" both often occur with "eat;" Unger et al., 2020; Wojcik & Saffran, 2015). Furthermore, toddlers draw inferences about the properties of novel nouns on the basis of the linguistic context they are embedded in (i.e., infer that "dax" refers to an animate object after hearing "the dax is crying;" Ferguson et al., 2014).

While language input clearly plays a key role in the development of children's word knowledge, it is not the sole source of information available during word learning moments. Children receive linguistic input within situated contexts; that is, children are engaging in particular actions with particular objects in particular places. For example, when a child hears "banana," they are likely to be in the kitchen or at the grocery store. Although environmental contexts like kitchens and grocery stores may make some aspects of word learning difficult (e.g., the fact that there are multiple objects present may make it more difficult to determine the correct referent for a given word), they may also serve as a valuable cue to aspects word meanings beyond simple object-label mappings. In particular, the physical locations where children encounter objects during word learning may provide important semantic information. Indeed, there are systematic relationships between the words children hear and the places where they hear them. Toddlers hear the highest proportion of food-related words in the kitchen and the highest proportion of body-related words in the bathroom (Custode & Tamis-LeMonda, 2020; Tamis-LeMonda et al., 2019). Thus, in children's naturalistic experiences, words often occur in particular locations that are related to their semantic properties, and therefore, these locations are potentially informative about word meanings.

Moreover, children are sensitive to these regularities in the locations where they typically encounter words and objects. Toddlers demonstrate an adult-like "scene-inconsistency" effect, recognizing when an object is semantically inconsistent with a particular scene (e.g., a banana in the bathroom; Helo et al., 2017; Maffongelli et al., 2020). In addition, toddlers can distinguish between familiar objects belonging to different contextual categories (e.g., kitchen objects vs. bathroom objects; Mandler et al., 1987). Children also encode information about scenes when learning about new objects and their labels. For example, after learning novel object-label mappings in a storybook, preschoolers remembered the scene in which each novel object appeared (Knabe & Vlach, 2022). In fact, scene-object associations were remembered more accurately than all other tested associations (e.g., person-object, scene-word, etc.). Thus, there is robust evidence that toddlers and preschoolers represent the environmental contexts within which objects are encountered.

The relationships between words, objects, and contexts also appear to influence children's word learning. Rich, naturalistic data from one child revealed that words the child heard in more consistent spatial contexts—for example, "moon," which was primarily heard in the bedroom—were learned earlier than words the child heard in more variable contexts (Roy et al., 2015). Follow-up work with largescale, cross-linguistic corpora reveals converging results: words used in more consistent conversational topic contexts (e.g., during mealtimes or getting dressed) are learned earlier (Unger et al., 2024). Furthermore, toddlers' early understanding of object labels is closely intertwined with the contexts in which they encounter them. For example, at 18 months, toddlers recognize a video of a person putting on shoes as a referent for the word "shoe," but not a video of a person performing a different action with shoes, like rubbing them together (Hagihara et al., 2022). These findings have led to growing interest in understanding the mechanisms underlying contextually-scaffolded word learning (Rowe & Weisleder, 2020; Tamis-LeMonda & Masek, 2023).

In-lab novel word learning paradigms have begun to directly manipulate environmental context to examine its impact on word learning success. There is evidence that 2- and 3-year-olds learn object labels and facts about objects better when they are presented in a consistent context across the training phase (Tippenhauer & Saylor, 2019; Vlach & Sandhofer, 2011). These studies instantiated context based on the color of the background the objects were presented on. However, the kinds of contexts in which children learn about object labels in everyday life are far more complex. This complexity may provide additional information that not only influences how well children encode object-label mappings, but also contributes to *what* children learn about novel words. For example, different contexts may activate different word learning biases. When toddlers were taught novel labels for novel non-solid objects, they were more likely to use a material bias to extend the labels to new exemplars when they learned about the objects while seated in a highchair compared to at a table (Perry et al., 2014). The authors suggest that toddlers' previous experience of frequently interacting with non-solid food objects while sitting in a highchair supported toddlers' material bias when learning new words in that same context. Thus, there is evidence that the contexts in which young children learn new words impacts how well they retain object-label mappings and what strategies they use to extend them. However, less is known regarding how environmental context might contribute to building other forms of word knowledge beyond the appropriate referents for a given label. Just as different contexts elicit different extension strategies, different contexts may license different inferences about the meanings of new words. In other words, can environmental context—specifically the types of scenes that objects are located in—help children situate new words within their existing semantic networks?

In the current study, we developed a word-learning paradigm that incorporated environmental context by presenting objects in different background scenes. During training, children observed four novel object-label pairs. Two objects were always pictured in kitchen scenes and two were always pictured in outdoor scenes. To test whether environmental contexts influenced children's understanding of the meanings of the object-label pairs, children completed a verb-mediated prediction task. The novel words were preceded either by a neutral verb (*find* or *see*) or a context-related verb (*eat* or *throw*). If children used environmental context to ascribe meaning to the novel words, then they should be able to predict the upcoming word based on the context-related verb. In other words, we predicted that children would anticipate that the label for an object previously presented in kitchen scenes would follow the verb "eat" while the label of an object presented in outdoor scenes would follow the verb "throw."

To demonstrate verb-mediated prediction, children must have robust verb comprehension and online language processing skills. Children demonstrate these skills by age three (Gambi et al., 2018; Meints et al., 2008), and therefore, we included three-year-olds (36 to 48 months) in our sample. Furthermore, the preschool years are an important stage of semantic development. By 3 years old, children are adept at recognizing semantic relationships between items that directly co-occur—often referred to a thematic associations (e.g. cat and mouse)—but find it more difficult to identify semantic relationships between items with shared co-occurrences—often referred to as taxonomic associations (e.g., cat and sheep; Unger et al., 2016, 2020; Unger & Fisher, 2021). Thus, environmental context may be a particularly valuable cue to help children at this age build more robust semantic networks and situate new words within them. If three-year-olds make verb-mediated predictions based on the context in which they encountered a novel object-label pair, this would suggest that environmental context does indeed affect children's semantic representations of new words. The study design, hypothesis, and analytic approach were pre-registered and can be accessed along with all stimuli, data, and analysis scripts on OSF (https://osf.io/35grb/).

2. Method

2.1. Participants

The sample size ($N = 46$) was pre-registered and identified with an a prior power analysis in G*Power to provide 90 % power based on effect sizes from previous research (Bobb et al., 2016; Lukyanenko & Fisher, 2016; Mani & Huettig, 2012). Thirty-two additional children were excluded from analysis due to data loss (e.g., inattention to the screen, see Supplementary Materials for more details). Participants were recruited using an existing database of interested families. All children were full-term (≤4 weeks early), were primarily English-learners (≤10 h/week of exposure to a language other than English), and had no hearing or vision concerns.

2.2. Materials

2.2.1. Novel objects and labels

The novel objects were created out of sculpting material to appear ambiguously toy-like and food-like (see Fig. 1a). The novel labels were selected from the NOUN Database (Horst & Hout, 2016), were disyllabic, and had similar English phonotactic probabilities (Vitevitch & Luce, 2004).

2.2.2. Training phase

The training phase introduced children to the four object-label pairs. Each object-label was yoked to a specific context (two in kitchen scenes and two in outdoor scenes). On each trial, children first saw a zoomed-out image of a scene with one of the novel objects embedded (see Fig. 1b). Then, the image zoomed into the novel object while it was labeled (e.g., "Look at this! This is a bosa! Wow, a bosa!"). Each object was labeled on 3 trials and embedded in 3 different images depicting its assigned context (i.e., objects in kitchen scenes appeared on the counter, in the fridge, and on a highchair tray; objects in outdoor scenes appeared in the sandbox, on the playground, and in the yard). The assignment of object-label pairs to kitchen versus outdoor scenes was counterbalanced across participants.

2.2.3. Test phase

Children completed a verb-mediated prediction task using a looking-while-listening paradigm (Fernald et al., 2008). Each trial included two images, an object trained in kitchen scenes and an object trained in outdoor scenes (see Fig. 1c). At test, both objects were pictured on blank gray backgrounds. While looking at these images, children heard a sentence containing the trained label for one of the objects. Labels were preceded by either a neutral verb ("I like to *find* the bosa") or a context-related verb. If the target label (e.g., "bosa") referred to an object that was trained in kitchen scenes, it would be preceded by the verb "eat" on context-related verb trials ("I like to *eat* the bosa"). If the target label referred to an object that was trained in outdoor scenes, it would be preceded by the verb "throw" on context-related verb trials ("I like to *throw* the bosa"). Test sentences were normalized for average intensity (70 dB) and duration (6000 ms). Trials began with the images in silence (1500 ms), then the carrier phrase "I like to" (800 ms), the verb (900 ms), the determiner "the" (450 ms), and finally the target word (850 ms), followed by 1500 ms of silence with the images still visible.

2.3. Procedure

Children sat alone or in their caregiver's lap in front of a 55" LCD screen in a soundproof booth. A video camera below the screen recorded children's eye movements for offline coding. The training phase

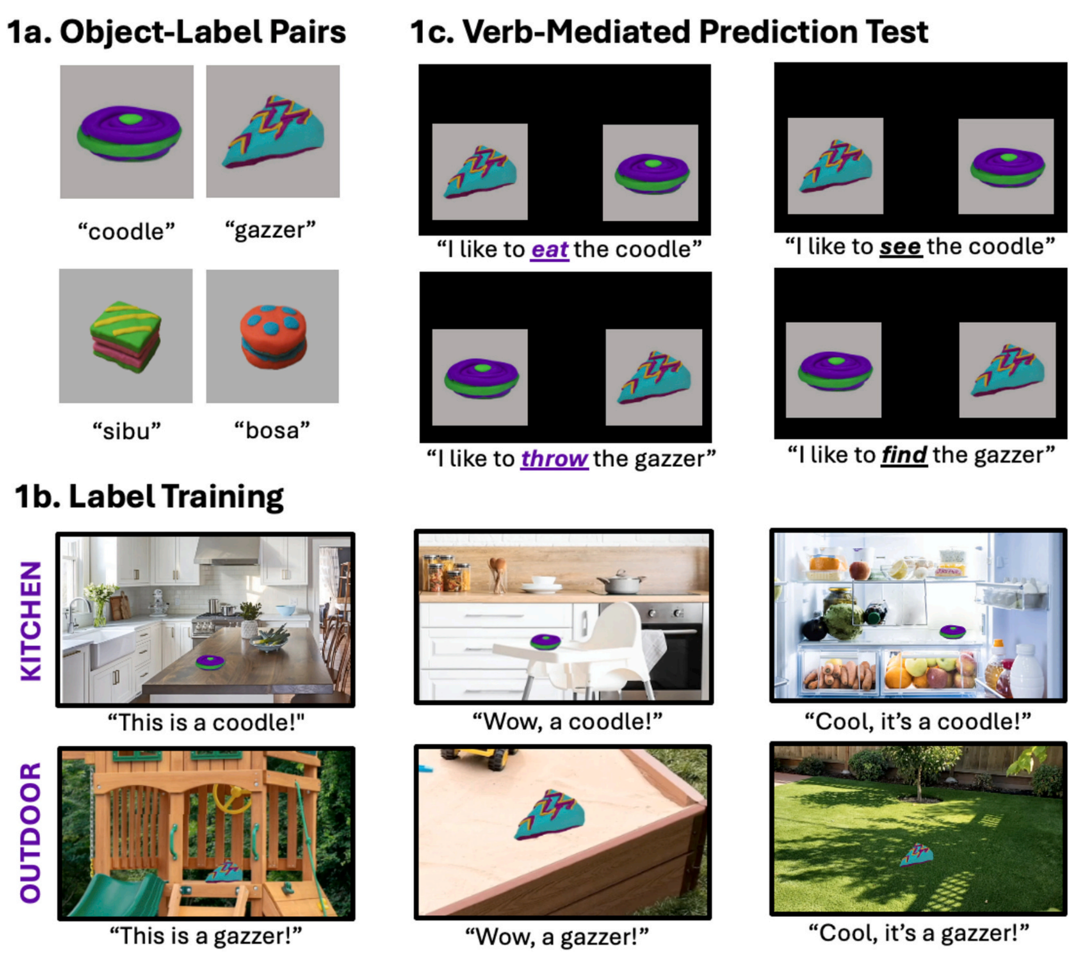

Fig. 1. The four novel object-label pairs used in the study (1a). Example label training trials for kitchen and outdoor scenes (1b). Example verb-mediated prediction test trials (1c).

consisted of 12 trials (3 trials for each of 4 novel words). The test phase consisted of 16 novel word trials. Two familiar word trials were included at the beginning of the test phase to introduce children to the task and two more were interspersed throughout the test to maintain interest. The 16 novel word trials included 8 Neutral Verb and 8 Context Verb trials. Each object was the target on 2 Neutral Verb and 2 Context Verb trials. Targets appeared on the left and right side of the screen equally often and the trial order was pseudorandomized such that the same target object was not tested on consecutive trials and no more than two trials of the same type (Neutral Verb and Context Verb) occurred consecutively. Attention-getters occurred after every 4 trials to maintain children's attention to the screen.

2.4. Measures

2.4.1. Gaze behavior

Eye gaze was manually coded frame-by-frame (33 ms) using Peyecoder (Olson et al., 2020). Gaze behavior was used to assess children's verb-mediated predictions and to measure their learning of the object-label pairs. Prediction behavior was measured in an anticipatory window spanning from 300 ms after verb onset until noun onset, which occurred 1350 ms after verb onset (Mani & Huettig, 2012). During this window, prediction was calculated based on the proportion of time children spent fixating the target object (i.e., the object labeled later in the sentence) out of the total time they spent fixating either the target object or the distractor object. If children incorporated environmental context into their representation of the novel object-label pairs, then they should increase their looking to the target object on Context Verb trials in this anticipatory window.

We also measured children's word learning by examining gaze behavior after they heard the noun in a target window from 300 ms to 1800 ms after noun onset (Fernald et al., 2008). Accuracy was calculated based on the proportion of time children fixated the target object out of the total time they fixated either the target or the distractor object. If children successfully learned the novel object-label mappings, they should continue to increase their looking to the target object in this target window.

Lastly, we measured reaction time. Reaction time was calculated as the time at which children first shifted their gaze to the target object after hearing the verb (Fernald et al., 2008; Hendrickson et al., 2017). This measure only included trials on which children were fixating the distractor object at verb onset. If children incorporated environmental context into their representation of the novel object-label pairs, they should shift their gaze to the target object earlier on Context Verb trials compared to Neutral Verb trials.

To be included in the analysis, children had to contribute at least 50 % usable frames (i.e., be attending to the screen for at least half of the 33 ms frames in both the anticipatory window and target window) for at least four test trials of each type (i.e., at least four Neutral Verb and at least four Context Verb trials). All videos were coded in silence by trained coders such that coders were unaware of which object was the target on each trial. To measure reliability, 25 % of the videos were independently recoded by a second coder. Coders agreed on the gaze location for 97.3 % of all frames and agreed within one frame on 92.9 % of all shifts in gaze.

2.4.2. Vocabulary survey

We collected a short, parent-report vocabulary measure to ensure that all participating children understood the verbs used in the study. If a child did not understand these verbs, they were excluded from analysis ($N = 1$).

3. Results

First, we measured whether children successfully learned the novel object-label mappings presented during training. We regressed accuracy (proportion of time looking at the target object out of the amount of time spent looking at either the target or distracter object) during the target window (300-1800 ms after target noun onset) on trial type (Context Verb and Neutral Verb). Accuracy was coded with an offset of −0.5 such that an accuracy value of 0 was equal to chance performance (i.e., looking equally often at the target and distracter). Trial type was centered with Context Verb trials coded as 0.5 and Neutral Verb trials coded as −0.5. The model with maximal random effect structure (i.e., a by-subject random intercept and by-subject random slope for trial type) did not converge. Following recommendations for resolving convergence errors, the covariance between the by-subject random intercept and by-subject random slope for condition was removed (Barr et al., 2013; Brauer & Curtin, 2018; Muradoglu et al., 2023). The results of this model confirmed that, on average, children looked towards the target object after hearing its label significantly more than expected by chance, $b = 0.15$, $F(1,44.03) = 74.95$, $p < .001$. Furthermore, there was not a significant difference in target window accuracy between the two trial types, $b = -0.003$, $F(1,43.17) = 0.01$, $p = .92$, suggesting that children successfully looked towards the correct object after hearing its label regardless of whether a context-related or neutral verb had come before it (see Fig. 2). These data confirm that children successfully learned the object-label mappings.

We then turned to our main question: did children acquire additional semantic knowledge based on the context in which they encountered the object-label pairs during training? This analysis focused on anticipation—looking to the target object *before* hearing its label—and compared looking behavior after a context-related verb (i.e., "eat" or "throw") or a neutral verb ("find" or "see"). Our pre-registered linear mixed-effects model analysis—regressing target looking (the proportion of time spent looking at the target object divided by the total time spent looking at either the target or distracter object) during the anticipatory window (300-1350 ms after verb onset) on trial type (Context Verb and Neutral Verb)—did not converge. Convergence errors could only be resolved by removing the by-subject random slope for trial type, yielding a model that should be interpreted with caution given that trial type was manipulated within subjects (Barr et al., 2013; Brauer & Curtin, 2018). Therefore, the results of this model are not reported in the main text, but can be found in the Supplementary Materials.

To address the difficulties with analyzing children's gaze behavior during the anticipatory window using the pre-registered linear mixed-effects model approach, we carried out a linear model analysis using a difference score. The difference score calculation accounts for the within-subjects nature of the data and therefore the model does not require random effects structure, avoiding the converge issues that arose from the linear mixed-effects model approach. We calculated each participant's average proportion of target looking during the anticipatory window for Context Verb trials and for Neutral Verb trials. Then, we subtracted their average anticipatory target looking on Neutral Verb trials from their average anticipatory target looking on Context Verb trials (Context Verb target looking - Neutral Verb target looking), resulting in a single difference score for each participant. Therefore, a difference score above zero indicated that a participant looked towards the target during the anticipatory window more on Context Verb trials than on Neutral Verb trials. We estimated an intercept-only linear model with participants' difference scores. The intercept of this model was significant, $b = 0.06$, $F(1,45) = 4.26$, $p = .045$, suggesting that, on average, participants' difference scores were significantly above 0. Therefore, participants looked towards the target object during the anticipatory window significantly more on Context Verb trials ($M = 0.54$, $SD = 0.36$) than on Neutral Verb trials ($M = 0.48$, $SD = 0.39$; see Fig. 2b). Consistent with our hypothesis, this result suggests that children encoded context during word learning and incorporated it into their semantic representations of the novel object-label mappings, allowing them to make predictions from context-related verbs.

In addition to examining children's target looking during the anticipatory window as a measure of prediction, we measured reaction time

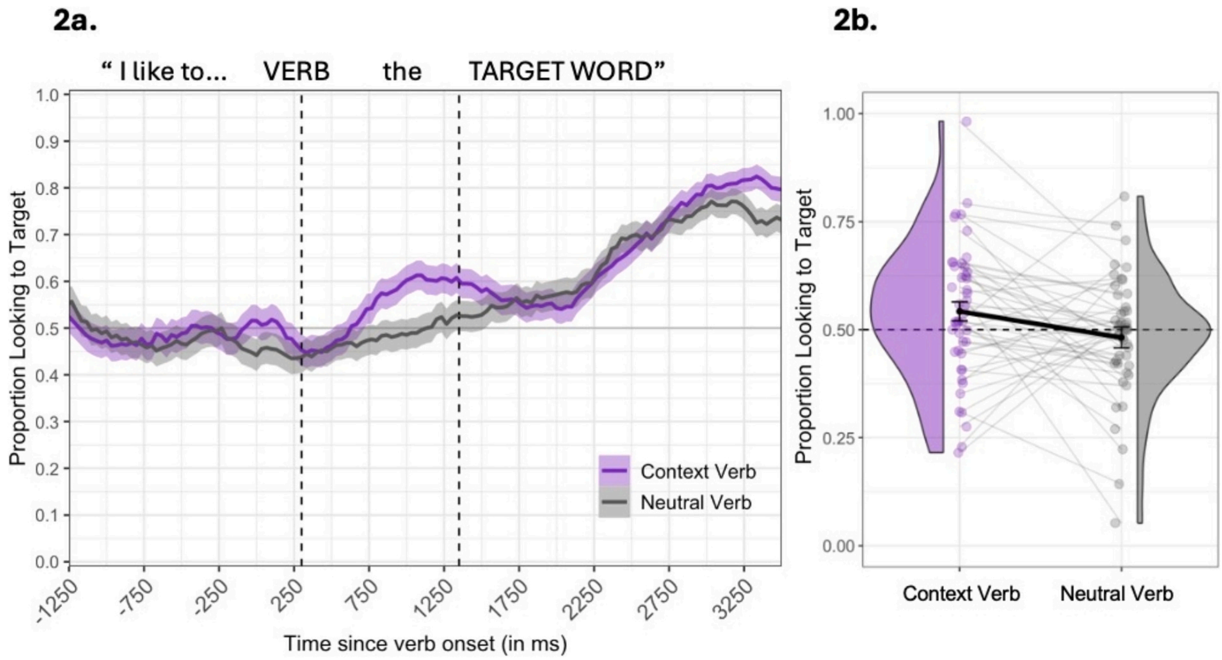

Fig. 2. **a.** Time course of children's gaze behavior on the verb-mediated prediction trials with Context Verbs in purple and Neutral Verbs in gray. Shaded area around the lines represents the standard error of the point estimates. Dashed vertical lines indicate the start and end of the anticipatory window. **b.** Proportion looking to target, averaged across the anticipatory window on Context Verb vs. Neutral Verb trials. Colored points indicate participant-level averages, black points indicate group-level averages for each trial type. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(the time at which children shifted their gaze off the distracter object and towards the target object). We estimated a linear mixed-effects model regressing reaction time on trial type (Context Verb and Neutral Verb). All participants who contributed data to the proportion looking analysis ($N = 46$) also contributed data to the reaction time analysis. However, only trials on which the child was looking at the distracter object at verb onset and later shifted to the target object were included in this analysis ($N = 184$ trials out of 522 possible; 91 Context Verb trials and 93 Neutral Verb trials; Fernald et al., 2008). Trial type was dummy coded (Context Verb trials 0 and Neutral Verb trials 1). This model did not converge with the maximal random effect structure, so the covariance between the by-subject random intercept and the by-subject random slope for trial type was removed. This model revealed a significant effect of trial type, $b = 229.54$, $F(1,79.4) = 4.42$, $p = .039$. Children shifted their gaze from the distracter to the target earlier on Context Verb trials ($M = 971$ ms, $SD = 605$) than on Neutral Verb trials ($M = 1210$ ms, $SD = 765$). This result provides further evidence that children were able to predict which object would be labeled based on the context-related verbs, indicating that they encoded contextual information during word learning. It is valuable to note that the average time at which children shifted their gaze to the target object on context-related verb trials (971 ms) was after verb offset (900 ms), indicating that children may have needed to hear most if not all of the verb to make a prediction about which object would likely be named next. An exploratory cluster-based permutation analysis, which can be found in the Supplementary Materials, lends further support to this possibility. As such, it may be important to consider the most appropriate window of analysis for verb-mediated prediction paradigms in future work (see Supplementary Materials for further detail).

4. Discussion

In children's everyday lives, they encounter words and objects in systematic environments (Custode & Tamis-LeMonda, 2020). The current study examined whether children use information about environmental context to situate novel object-label pairs within their existing semantic network. Three-year-olds learned labels for objects that were pictured either in kitchen scenes or outdoor scenes. Then, to measure whether children incorporated environmental context into their representations of the novel object-label pairs, children completed a verb-mediated prediction paradigm where each novel noun was preceded either by a neutral verb ("find" or "see") or a context-related verb ("eat" or "throw"). Specifically, we measured whether children predicted that the label for an object previously pictured in kitchen scenes would follow the verb "eat" and the label for an object previously pictured in outdoor scenes would follow the verb "throw." Children did make these predictions, looking to the target object significantly more accurately and quickly on context-related verb trials compared to neutral verb trials. Put another way, children were able to anticipate the target noun based on verb semantics, despite having never heard the object-label pairs used with the verbs before. Moreover, these aspects of meaning were not directly demonstrated—children never saw the objects being eaten or thrown. This pattern of results is consistent with the hypothesis that children encode environmental contexts during word learning and use their prior knowledge about those contexts to infer aspects of the meanings of novel words

To our knowledge, this study is the first to manipulate the scene context in which novel object-label mappings are embedded and to examine the impact of scene context on children's semantic representations of the object-label pairs. While previous research has demonstrated that children encode the scenes in which they encounter novel objects and their labels (e.g., Knabe & Vlach, 2022), the current study further illustrates that children can use their knowledge about these contexts to situate novel object-label pairs within their existing semantic network. Children's semantic representations of the novel object-label pairs were influenced by a simple visual cue to environmental context during learning: the background scenes in which the objects were embedded in during training. Although the visual cue to context was no longer available during the verb-mediated prediction trials, children nonetheless connected the object-label pairs to context-related verbs based on the contexts in which they originally saw the objects. Even in the absence of richer contextual cues that children have access to in their naturalistic experiences—for example, linguistic, temporal, and interactional cues (Alhama et al., 2023; Roy et al., 2015; Tamis-LeMonda et al., 2019)—visual context was sufficient to influence children's semantic representations of novel words. This result suggests that environmental contexts, even those as simple as visual scenes, can inform children's understanding of words.

In the current paradigm, it is difficult to disentangle exactly what kinds of semantic inferences might underlie children's behavior on the verb-mediated prediction task. For example, it could be the case that children looked to the object trained in kitchen scenes after hearing the verb "eat" because they inferred that it was an edible food. Alternatively, children could have looked to the object trained in kitchen scenes after hearing the verb "eat" due to their shared association with kitchens (i.e., they associated both the object and the verb "eat" with kitchens). Indeed, recent work has suggested that these kinds of associative matrices and shared co-occurrences likely play an important role in word learning and semantic development (Knabe & Vlach, 2022; Unger & Fisher, 2021; Wojcik & Saffran, 2015). Regardless of the specific inferences underlying children's predictions on the Context Verb trials, the fact that children could predict the upcoming noun on the basis of context-related verbs indicates that children must have encoded environmental context during word learning and subsequently used that information to situate the novel nouns within their existing semantic network.

It is interesting to note that these prediction effects did not extend into the target window. In other words, in the anticipatory window (i.e., after hearing the verb but before hearing the noun), children looked more to the target on Context Verb trials than on Neutral Verb trials; however, in the target window (i.e., after hearing the target word), children's looking to the target did not differ between our two trial types. Visual inspection of the time course plot (see Fig. 2) reveals that, on Context Verb trials, children appear to be shifting their gaze away from the target at the beginning of the target window, after having fixated on it in the anticipatory window. This behavior may have prevented our prediction effects from further facilitating word recognition in the target window. Indeed, this behavior is consistent with previous investigations of children's verb-mediated prediction with familiar words: prediction is evident in the anticipatory window, but no differences between semantically-constraining and neutral verb trials are found in the target window (Mani & Huettig, 2012). It is possible that shortening the anticipatory window (i.e., reducing the time between the verb and the noun) would reveal persisting effects in the target window; however, a shorter anticipatory window could also weaken predictive effects if children have less time to shift their gaze before hearing the target word.

This study adds to efforts to design in-lab novel word learning tasks that directly manipulate elements of environmental context in order to examine its casual impact on word learning. The current study focused on one particular aspect of environmental context—location, or scene context—while other paradigms have investigated additional elements of the environmental context including how objects interact with the physical environment around them and what other objects are present during word learning (e.g., Benitez & Smith, 2012; Breitfeld & Saffran, 2024; Chen & Yu, 2022; Luchkina & Waxman, 2021; Pomper & Saffran, 2019). Continuing to develop novel word learning paradigms that can examine the effects of contextual input on word learning outcomes may provide additional insights into early word learning trajectories. Though the results of this study suggest that, on average, children incorporated environmental context in their understanding of novel object-label pairs, there may be important individual differences in this process. Some children may be less attentive to and less likely to encode associations between objects and contexts. For example, children with Developmental Language Disorder are less sensitive to semantic object-scene inconsistencies compared to their peers following typical language developmental trajectories (Helo et al., 2022). Furthermore, different children might have different associations with a given context. One child might exclusively engage in food-related activities in the kitchen, while another might participate in lots of different activities in the kitchen—eating food, playing with toys, etc. Therefore, the representation one child forms about a novel object-label pair encountered in the kitchen may be very different from the representation another child forms. Cultural differences in the kinds of activities engaged in and the kinds of language used within different contexts may also influence these associations. For example, children in different communities may interact with different types of objects (i.e., toys vs. natural objects) and transition between interacting with different unique objects and object categories more or less frequently (Casey et al., 2022; Herzberg et al., 2022). The current study is limited to the experiences of a largely White sample of children living in the midwestern United States, which is likely unrepresentative of children's experiences in other cultures.

5. Conclusion

In sum, the current study suggests that children encode environmental context during word learning and use it to situate novel object-label pairs into their existing semantic network. Furthermore, our results demonstrate that children were able to use this context-based semantic knowledge in the service of online language prediction. Even in a simple paradigm where environmental context was instantiated as the background content of an image, children incorporated this context into their novel word representations, building meaning beyond simple object-label mappings. These results underscore the importance of environmental context in early world learning and encourage further investigation of how children acquire rich, multifaceted word knowledge.

Data statement

The analyses presented here were preregistered. The data and analytic code necessary to reproduce the analyses presented in this paper are publicly accessible. The materials necessary to attempt to replicate the findings presented here are also publicly accessible. The data that support the findings of this study are openly available in the Open Science Framework (OSF) at https://osf.io/35grb/.

CRediT authorship contribution statement

Elise Breitfeld: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jenny R. Saffran:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors confirm that there are no interests or relationships, financial or otherwise, that might be perceived as a potential source of conflict of interest.

Acknowledgements

This work was supported by grants from NICHD awarded to EB (F31HD115328), JRS (R01HD105313), and the Waisman Center (U54HD090256). We would like to thank the participating families and members of the Infant Learning Lab. In particular, we thank Anna Compton, Emily Kassens, and Dasha Yermol for their help with data collection, as well as Sydney Lex, Rayane Prado Nunes, Julia Waltzman, and Evann Zima for their help with data coding. Special thanks to Martha Alibali, Desia Bacon, Jess Mankewitz, Ron Pomper, and Haley Weaver for their valuable input and discussion throughout the project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2025.106162.

Data availability

All data is available on the Open Science Framework (OSF) at https://osf.io/35grb/.

References

Alhama, R. G., Rowland, C. F., & Kidd, E. (2023). How does linguistic context influence word learning? *Journal of Child Language, 50*(6), 1374–1393. https://doi.org/10.1017/S0305000923000302

Arias-Trejo, N., & Plunkett, K. (2013). What's in a link: Associative and taxonomic priming effects in the infant lexicon. *Cognition, 128*(2), 214–227. https://doi.org/10.1016/j.cognition.2013.03.008

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3). https://doi.org/10.1016/j.jml.2012.11.001

Benitez, V. L., & Smith, L. B. (2012). Predictable locations aid early object name learning. *Cognition*. https://doi.org/10.1016/j.cognition.2012.08.006

Bobb, S. C., Huettig, F., & Mani, N. (2016). Predicting visual information during sentence processing: Toddlers activate an object's shape before it is mentioned. *Journal of Experimental Child Psychology, 151*, 51–64. https://doi.org/10.1016/j.jecp.2015.11.002

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of non-independent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods, 23*(3), 389–411. https://doi.org/10.1037/met0000159

Breitfeld, E., & Saffran, J. R. (2024). Early word learning is influenced by physical environments. *Child Development, 95*(3), 962–971. https://doi.org/10.1111/cdev.14046

Casey, K., Elliott, M., Mickiewicz, E., Silva Mandujano, A., Shorter, K., Duquette, M., … Casillas, M. (2022). Sticks, leaves, buckets, and bowls: Distributional patterns of children's at-home object handling in two subsistence societies. *Proceedings of the Annual Meeting of the Cognitive Science Society, 44*(44). https://escholarship.org/uc/item/6wx2x30s.

Chen, C., & Yu, C. (2022). Building lexical networks: Preschoolers extract different types of information in cross-situational learning. *Journal of Experimental Child Psychology, 220*, Article 105430. https://doi.org/10.1016/j.jecp.2022.105430

Custode, S. A., & Tamis-LeMonda, C. (2020). Cracking the code: Social and contextual cues to language input in the home environment. *Infancy, 25*(6), 809–826. https://doi.org/10.1111/infa.12361

Ferguson, B., Graf, E., & Waxman, S. R. (2014). Infants use known verbs to learn novel nouns: Evidence from 15- and 19-month-olds. *Cognition, 131*(1), 139–146. https://doi.org/10.1016/j.cognition.2013.12.014

Fernald, A., Marchman, V. A., Zangl, R., & Portillo, A. L. (2008). *Using Eye Movements to Monitor Spoken Language Comprehension by Infants and Young Children*. Looking While Listening.

Gambi, C., Gorrie, F., Pickering, M. J., & Rabagliati, H. (2018). The development of linguistic prediction: Predictions of sound and meaning in 2- to 5-year-olds. *Journal of Experimental Child Psychology, 173*, 351–370. https://doi.org/10.1016/j.jecp.2018.04.012

Gelman, S. A., & Raman, L. (2003). Preschool children use linguistic form class and pragmatic cues to interpret generics. *Child Development, 74*(1), 308–325. https://doi.org/10.1111/1467-8624.00537

Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cognitive Psychology, 61*(3), 273–301. https://doi.org/10.1016/j.cogpsych.2010.06.001

Gelman, S. A., Ware, E. A., Kleinberg, F., Manczak, E. M., & Stilwell, S. M. (2014). Individual differences in children's and parents' generic language. *Child Development*. https://doi.org/10.1111/cdev.12187

Graham, S. A., Nayer, S. L., & Gelman, S. A. (2011). Two-year-olds use the generic/non-generic distinction to guide their inferences about novel kinds. *Child Development, 82*(2), 493–507. https://doi.org/10.1111/j.1467-8624.2010.01572.x

Hagihara, H., Yamamoto, H., Moriguchi, Y., & Sakagami, M. (2022). When "shoe" becomes free from "putting on": The link between early meanings of object words and object-specific actions. *Cognition, 226*, Article 105177. https://doi.org/10.1016/j.cognition.2022.105177

Helo, A., Azaiez, N., & Rämä, P. (2017). Word processing in scene context: An event-elated potential study in young children. *Developmental Neuropsychology, 42*(7–8). https://doi.org/10.1080/87565641.2017.1396604

Helo, A., Guerra, E., Coloma, C. J., Aravena-Bravo, P., & Rämä, P. (2022). Do children with developmental language disorder activate scene knowledge to guide visual attention? Effect of object-scene inconsistencies on gaze allocation. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.796459

Hendrickson, K., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2017). Assessing a continuum of lexical–semantic knowledge in the second year of life: A multimodal approach. *Journal of Experimental Child Psychology, 158*, 95–111. https://doi.org/10.1016/j.jecp.2017.01.003

Herzberg, O., Fletcher, K. K., Schatz, J. L., Adolph, K. E., & Tamis-LeMonda, C. S. (2022). Infant exuberant object play at home: Immense amounts of time-distributed, variable practice. *Child Development, 93*(1), 150–164. https://doi.org/10.1111/cdev.13669

Horst, J. S., & Hout, M. C. (2016). The novel object and unusual name (NOUN) database: A collection of novel images for use in experimental research. *Behavior Research Methods*. https://doi.org/10.3758/s13428-015-0647-3

Knabe, M. L., & Vlach, H. A. (2022). Not all is forgotten: Children's associative matrices for features of a word learning episode. *Developmental Science*. , Article e13291. https://doi.org/10.1111/desc.13291

Luchkina, E., & Waxman, S. R. (2021). Semantic priming supports infants' ability to learn names of unseen objects. *PLoS ONE, 16*(1), Article e0244968. https://doi.org/10.1371/journal.pone.0244968

Lukyanenko, C., & Fisher, C. (2016). Where are the cookies? Two- and three-year-olds use number-marked verbs to anticipate upcoming nouns. *Cognition, 146*, 349–370. https://doi.org/10.1016/j.cognition.2015.10.012

Maffongelli, L., Öhlschläger, S., & Lê-Hoa Võ, M. (2020). The development of scene semantics: First ERP indications for the processing of semantic object-scene inconsistencies in 24-month-olds. *Collabra Psychology, 6*(1), Article 17707. https://doi.org/10.1525/collabra.17707

Mandler, J. M., Fivush, R., & Reznick, J. S. (1987). The development of contextual categories. *Cognitive Development, 2*(4), 339–354. https://doi.org/10.1016/S0885-2014(87)80012-6

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake-but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4). https://doi.org/10.1037/a0029284

Meints, K., Plunkett, K., & Harris, P. (2008). Eating apples and houseplants: Typicality constraints on thematic roles in early verb learning. *Language & Cognitive Processes, 23*, 434–463. https://doi.org/10.1080/01690960701726232

Muradoglu, M., Cimpian, J. R., & Cimpian, A. (2023). Mixed-effects models for cognitive development researchers. *Journal of Cognition and Development, 1–34*. https://doi.org/10.1080/15248372.2023.2176856

Olson, R. H., Pomper, R., Potter, C. E., Hay, J. F., Saffran, J. R., Ellis Weismer, S., & Lew-Williams, C. (2020). *Peyecoder: An Open-Source Program for Coding Eye Movements [Computer Software]*. Zenodo. https://doi.org/10.5281/zenodo.3939234

Pappas, A., & Gelman, S. A. (1998). Generic noun phrases in mother–child conversations. *Journal of Child Language, 25*(1), 19–33. https://doi.org/10.1017/S0305000997003292

Perry, L. K., Samuelson, L. K., & Burdinie, J. B. (2014). Highchair philosophers: The impact of seating context-dependent exploration on children's naming biases. *Developmental Science*. https://doi.org/10.1111/desc.12147

Pomper, R., & Saffran, J. R. (2019). Familiar object salience affects novel word learning. *Child Development*. https://doi.org/10.1111/cdev.13053

Quine, W. V. O. (1960). *Word and Object* (New ed.). MIT Press.

Rowe, M. L., & Weisleder, A. (2020). Language development in ontext. *Annual Review of Developmental Psychology, 2*(1), 201–223. https://doi.org/10.1146/annurev-devpsych-042220-121816

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences of the United States of America, 112*(41). https://doi.org/10.1073/pnas.1419773112

Tamis-LeMonda, C. S., Custode, S., Kuchirko, Y., Escobar, K., & Lo, T. (2019). Routine language: Speech irected to infants during home activities. *Child Development, 90*(6). https://doi.org/10.1111/cdev.13089

Tamis-LeMonda, C. S., & Masek, L. R. (2023). Embodied and embedded learning: Child, caregiver, and context. *Current Directions in Psychological Science, 32*(5), 369–378. https://doi.org/10.1177/09637214231178731

Tippenhauer, N., & Saylor, M. M. (2019). Effects of context variability on 2-year-olds' fact and word learning. *Journal of Experimental Child Psychology, 179*, 248–259. https://doi.org/10.1016/j.jecp.2018.11.011

Unger, L., Chang, T., Savic, O., Bergen, B. K., & Sloutsky, V. M. (2024). When is a word in good company for learning? *Developmental Science, 27*(5), Article e13510. https://doi.org/10.1111/desc.13510

Unger, L., & Fisher, A. V. (2021). The emergence of richly organized semantic knowledge from simple statistics: A synthetic review. *Developmental Review, 60*, Article 100949. https://doi.org/10.1016/j.dr.2021.100949

Unger, L., Fisher, A. V., Nugent, R., Ventura, S. L., & MacLellan, C. J. (2016). Developmental changes in semantic knowledge organization. *Journal of Experimental Child Psychology, 146*, 202–222. https://doi.org/10.1016/j.jecp.2016.01.005

Unger, L., Vales, C., & Fisher, A. V. (2020). The role of co-occurrence statistics in developing semantic knowledge. *Cognitive Science, 44*(9), Article e12894. https://doi.org/10.1111/cogs.12894

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers, 36*(3), 481–487.

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology*. https://doi.org/10.1016/j.jecp.2010.09.011

Wei, R., Kirby, A., Naigles, L. R., & Rowe, M. L. (2022). Parents' talk about conceptual categories with infants: Stability, variability, and implications for expressive language development. *Journal of Child Language, 1–22*. https://doi.org/10.1017/S0305000922000319

Willits, J. A., Wojcik, E. H., Seidenberg, M. S., & Saffran, J. R. (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: Evidence from auditory priming. *Infancy, 18*(6), 1053–1075.

Wilson, K., Frank, M. C., & Fourtassi, A. (2023). Conceptual hierarchy in child-directed speech: Implicit cues are more reliable. *Journal of Cognition and Development, 24*(4), 563–580. https://doi.org/10.1080/15248372.2023.2178436

Wojcik, E. H., & Saffran, J. R. (2015). Toddlers encode similarities among novel words from meaningful sentences. *Cognition, 138*, 10–20. https://doi.org/10.1016/j.cognition.2015.01.015

Wojcik, E. H., Zettersten, M., & Benitez, V. L. (2022). The map trap: Why and how word learning research should move beyond mapping. *WIREs Cognitive Science, 13*(4), Article e1596. https://doi.org/10.1002/wcs.1596